IMA Journal of Numerical Analysis (1994) **14**, 363–379

On the error of general linear methods for stiff dissipative differential equations

WILLEM HUNDSDORFER
Université de Genève, Section de Mathématiques, Case Postale 240, CH-1211 Genève 24, Switzerland

[Received 9 December 1991 and in revised form 31 May 1993]

Many numerical methods to solve initial value problems of the form $y' = f(t, y)$ can be written as general linear methods. Classical convergence results for such methods are based on the Lipschitz constant and bounds for certain partial derivatives of f. For stiff problems these quantities may be very large, and consequently the classical order of convergence loses its significance. In this paper we consider bounds for the global errors which are based only on bounds for derivatives of y for linear and non-linear dissipative problems with arbitrary stiffness.

1. Introduction

Linear multistep, Nordsieck and Runge–Kutta methods are popular methods for solving stiff initial value problems. Such methods, and combinations thereof, can be written as general linear methods.

If the initial value problem is given by

$$y'(t) = f(t, y(t)) \tag{1.1}$$

with known $y(0)$, a general linear method can be formulated as

$$u_{n+1,i} = \sum_{j=1}^{k} a_{ij}^{(1)} u_{nj} + h \sum_{j=1}^{s} b_{ij}^{(1)} f(t_n + c_j h, v_{nj}), \qquad i = 1, 2, \ldots, k, \tag{1.2a}$$

$$v_{ni} = \sum_{j=1}^{k} a_{ij}^{(2)} u_{nj} + h \sum_{j=1}^{s} b_{ij}^{(2)} f(t_n + c_j h, v_{nj}), \qquad i = 1, 2, \ldots, s \tag{1.2b}$$

with stepsize $h > 0$ and $t_n = nh$, $n \geqslant 0$. The vectors v_{ni} are internal quantities and can be considered as within a black-box, whereas the u_{ni} are the relevant (external) results approximating given correct value functions $u_i(t_n)$. These $u_i(t_n)$ can stand for function values $y(t_n + \theta_i h)$ of the exact solution, scaled derivatives $hy'(t_n + \theta_i h)$, $h^2 y''(t_n + \theta_i h), \ldots$ or linear combinations of such terms. For a comprehensive discussion of general linear methods we refer the reader to the books by Butcher (1987b), Hairer *et al* (1987), Hairer & Wanner (1991) and the review paper of Burrage & Chipman (1989).

Classical error bounds for method (1.2) involve the Lipschitz constant of f and bounds for certain partial derivatives of $f(t, y)$ at $y = y(t)$, $t \geqslant 0$. For stiff problems these quantities can become very large, and therefore the error bounds

† Present address: CWI, P.O. Box 4079, 1009 AB Amsterdam, The Netherlands.

based on them become meaningless. In this paper we shall consider for two classes of well known stiff model problems error bounds in which only derivatives of $y(t)$ appear. This allows arbitrary stiffness.

The first class contains the linear equations

$$y'(t) = Ly(t) + g(t) \tag{1.3}$$

with a matrix L satisfying

$$\langle v, Lv \rangle \leq 0 \tag{1.4}$$

for all vectors v, with respect to the Euclidian inner product $\langle v, w \rangle = v^T w$. The second class consists of all non-linear dissipative problems (1.1), where f satisfies

$$\langle f(t, v) - f(t, w), v - w \rangle \leq 0 \tag{1.5}$$

for arbitrary vectors v, w.

It will always be assumed that the solution $y(t)$ is sufficiently smooth and that exact starting values $u_{0i} = u_i(t_0)$ are given. Under appropriate stability assumptions one can easily obtain a stiffness-independent convergence result, involving only derivatives of $y(t)$, with an order of convergence equal to the stage order q, see Section 2. Convergence results of this type are well known for Runge–Kutta methods, see for instance Dekker & Verwer (1984), Frank *et al* (1985b) and Hairer & Wanner (1991). In the latter reference such results were also derived for multistep Runge–Kutta methods.

The central issue in this paper is whether these convergence results with order q are optimal. In Section 3 it will be shown that for the linear problems one often has order $q + 1$, as is the case for Runge–Kutta methods, see Burrage *et al* (1986), Dekker *et al* (1986). In Section 4 we will prove that for the class of non-linear dissipative problems convergence with order $q + 1$ is only possible for methods where the stage order can be increased by changing the correct value function (i.e., changing the interpretation of the method). This generalizes a result of Burrage & Hundsdorfer (1987) for Runge–Kutta methods.

2. Preliminaries

Let $A_1 = (a_{ij}^{(1)})$, $A_2 = (a_{ij}^{(2)})$, $B_1 = (b_{ij}^{(1)})$, $B_2 = (b_{ij}^{(2)})$ and

$$u_n = (u_{n1}, \ldots, u_{nk})^T, \qquad v_n = (v_{n1}, \ldots, v_{ns})^T,$$

$$F_n(v_n) = \big(f(t_n + c_1 h, v_{n1}), \ldots, f(t_n + c_s h, v_{ns})\big)^T.$$

The vector u_n approximates $u(t_n) = \big(u_1(t_n), \ldots, u_k(t_n)\big)^T$. Method (1.2) can be written in compact form as

$$u_{n+1} = A_1 u_n + hB_1 F_n(v_n), \tag{2.1a}$$

$$v_n = A_2 u_n + hB_2 F_n(v_n). \tag{2.1b}$$

We shall use this notation also if f is vector valued, i.e. A_m, B_m are also used to denote the Kronecker products $A_m \otimes I$, $B_m \otimes I$, $m = 1, 2$, with I the identity matrix with the dimension of f.

Throughout this paper we shall work with the Euclidian inner product $\langle v, w\rangle = v^*w$ and norm $\|v\| = \langle v, v\rangle^{\frac{1}{2}}$ for real and complex vectors. The induced spectral norm for matrices is also denoted by $\|\cdot\|$.

A general starting point for a stability/convergence analysis is to consider along with (2.1) a perturbed scheme

$$\tilde{u}_{n+1} = A_1\tilde{u}_n + hB_1F_n(\tilde{v}_n) + \xi_n, \tag{2.2a}$$

$$\tilde{v}_n = A_2\tilde{u}_n + hB_2F_n(\tilde{v}_n) + \eta_n \tag{2.2b}$$

with perturbations ξ_n, η_n and starting error $\tilde{u}_0 - u_0$. Let

$$Z_n = \operatorname{diag}(z_{n1},\dots, z_{ns})$$

by the block-diagonal matrix defined by

$$Z_n = h\int_0^1 F_n'(v_n + \theta(\tilde{v}_n - v_n))\,\mathrm{d}\theta, \tag{2.3}$$

so that $Z_n(\tilde{v}_n - v_n) = hF_n(\tilde{v}_n) - hF_n(v_n)$. By subtraction of (2.1) from (2.2) is follows easily that

$$\tilde{v}_n - v_n = (I - B_2Z_n)^{-1}(A_2(\tilde{u}_n - u_n) + \eta_n) \tag{2.4}$$

and

$$\tilde{u}_{n+1} - u_{n+1} = R(Z_n)(\tilde{u}_n - u_n) + \xi_n + B_1Z_n(I - B_2Z_n)^{-1}\eta_n \tag{2.5}$$

where

$$R(Z_n) = A_1 + B_1Z_n(I - B_2Z_n)^{-1}A_2. \tag{2.6}$$

So, the matrix $R(Z_n)$ determines how the previous error $\tilde{u}_n - u_n$ will affect $\tilde{u}_{n+1} - u_{n+1}$, whereas the influence of the local perturbation η_n is governed by $B_1Z_n(I - B_2Z_n)^{-1}$.

We want to ensure that a small initial error and small perturbations cause only small global errors. This leads to stability requirements of the type

$$\|R(Z_n)R(Z_{n-1})\cdots R(Z_m)\| \leq \alpha, \tag{2.7}$$

$$\|B_1Z_n(I - B_2Z_n)^{-1}\| \leq \beta \tag{2.8}$$

for all $h > 0$ and $1 \leq m \leq n \leq N$, with constants α, β and with N the total number of steps needed to cover the integration interval at stepsize h. Whether these conditions are fulfillled will depend of course on the method and the class of problems under consideration. In the next sections this will be specified for the two classes of model problems.

Condition (2.7) represents the usual step-by-step stability. In general it holds that (2.7)$\Rightarrow$(2.8), but there are exceptions to this rule, for instance with Runge–Kutta methods of the Lobatto IIIC type on the class of non-linear dissipative problems, see Spijker (1986). To ensure that the system of algebraic equations (2.1*b*) is well conditioned one can require in addition to (2.8) that

$$\|(I - B_2Z_n)^{-1}\| \leq \beta. \tag{2.9}$$

Again this in general is satisfied if (2.7) holds.

To obtain bounds on the global discretization errors $u(t_n) - u_n$ which depend only on derivatives of $y(t)$, we set in formula (2.2)

$$\bar{u}_n = u(t_n), \qquad \bar{v}_n = v(t_n) := \big(y(t_n + c_1 h), \ldots, y(t_n + c_s h)\big)^{\mathrm{T}}$$

and use Taylor series expansions

$$u(t_n) = k_0 y(t_n) + k_1 h y'(t_n) + k_2 h^2 y''(t_n) + \cdots,$$
$$v(t_n) = e y(t_n) + c h y'(t_n) + \tfrac{1}{2} c^2 h^2 y''(t_n) + \cdots$$

with k-dimensional vectors k_j determined by the definition of the correct value function, and $e = (1, \ldots, 1)^{\mathrm{T}}$, $c^j = (c_1^j, \ldots, c_s^j)^{\mathrm{T}}$. Again, if $y(t)$ is vector valued we use k_j, e, c^j also to denote their Kronecker products with I. With this choice for $\bar{u}_n$, $\bar{v}_n$ we have

$$\xi_n = u(t_{n+1}) - A_1 u(t_n) - hB_1 v'(t_n), \qquad \eta_n = v(t_n) - A_2 u(t_n) - hB_2 v'(t_n),$$

and so we easily obtain expansions of the form

$$\begin{pmatrix} \xi_n \\ \eta_n \end{pmatrix} = \begin{pmatrix} d_1^{(q+1)} \\ d_2^{(q+1)} \end{pmatrix} h^{q+1} y^{(q+1)}(t_n) + \begin{pmatrix} d_1^{(q+2)} \\ d_2^{(q+2)} \end{pmatrix} h^{q+2} y^{(q+2)}(t_n) + \cdots \tag{2.10}$$

where the integer q (the stage order) and the vectors $d_1^{(j)}$, $d_2^{(j)}$ are determined by the method together with the definition of the correct value function. To have stage order ≥ 0 it should hold that $A_1 k_0 = k_0$, $A_2 k_0 = e$. This will always be assumed. The vector k_0 is called the preconsistency vector.

Example Suppose that $u_i(t_n) = y(t_n + \theta_i h)$. Let $\theta = (\theta_1, \ldots, \theta_k)^{\mathrm{T}}$. Then some calculations give an expansion (2.10) with

$$d_1^{(j)} = \frac{1}{j!}\big((\theta + e)^j - A_1 \theta^j - jB_1 c^{j-1}\big), \qquad d_2^{(j)} = \frac{1}{j!}(c^j - A_2 \theta^j - jB_2 c^{j-1})$$

for $j = 0, 1, \ldots$. Thus q is the largest integer such that $d_1^{(j)} = 0$, $d_2^{(j)} = 0$, for $j = 0, 1, \ldots, q$.

The expansions leading to (2.10) can be truncated at any level of h^j with remainder term involving only method-depending constants and $y^{(j)}$ evaluated in some intermediate points. For example, if $y \in C^{q+1}$ then ξ_n and η_n can be bounded in norm by $C_0 h^{q+1} \max_t \|y^{(q+1)}(t)\|$ with C_0 determined by the coefficients of the method.

3. Linear problems

3.1 *Linear scalar problems*

Before dealing with linear systems, we first consider the scalar case

$$y'(t) = \lambda y(t) + g(t) \tag{3.1}$$

with $\lambda \in \mathbb{C}$, $\operatorname{Re} \lambda \leq 0$. From (2.5), (2.6) we see that the global errors $\varepsilon_n = u(t_n) - u_n$ satisfy

$$\varepsilon_{n+1} = R(z)\varepsilon_n + \delta_n \tag{3.2}$$

with $z = h\lambda$, $R(z) = A_1 + zB_1(I - zB_2)^{-1}A_2$ and local errors δ_n given by

$$\delta_n = \sum_{j \geq q+1} \left(d_1^{(j)} + zB_1(I - zB_2)^{-1}d_2^{(j)}\right)h^j y^{(j)}(t_n). \tag{3.3}$$

We shall consider convergence uniformly for $z = h\lambda$ in some given subset D of the stability region

$$S = \{z \in \bar{\mathbb{C}} : R(z) \text{ power bounded}\}.$$

Then the stability conditions (2.7), (2.8) read

$$\sup_{n \geq 0} \|R(z)^n\| \leq \alpha, \qquad z \in D, \tag{3.4}$$

$$\|zB_1(I - zB_2)^{-1}\| \leq \beta, \qquad z \in D. \tag{3.5}$$

We assume for the moment that this holds and that $I - zB_2$ is non-singular on D. In the next subsection we shall take a closer look at these conditions. Further it will be tacitly assumed throughout the rest of the paper that $\det\left(I - R(z)\right)$ is not identically equal to 0. Then $\rho\left(R(z)\right) < 1$ in the interior of S, with ρ denoting the spectral radius.

Let D be fixed, so that α and β can be considered as constants determined by the method. From (3.4), (3.5) we obtain in a straightforward way a global error bound $Ct_n h^q$ with a constant C depending only on y (in fact, depending only on $\max_t \|y^{(q+1)}(t)\|$). In the following it will be discussed whether such a result is possible with order $q + 1$. For this we consider the leading term in the local error, and denote for simplicity

$$d_1 = d_1^{(q+1)}, \qquad d_2 = d_2^{(q+1)}.$$

Let

$$\varphi(z) = \left(I - R(z)\right)^{-1}\left(d_1 + zB_1(I - zB_2)^{-1}d_2\right). \tag{3.6}$$

The components of this vector valued function are rational functions in z. For those values of z for which $I - R(z)$ is singular we define $\varphi(z)$ by taking the limit values of these functions. This gives $\|\varphi(z)\| = \infty$ if z is a non-removable singularity of a component of φ. As we shall see, the singularity at the origin usually is removable.

THEOREM 3.1 Assume (3.4) and (3.5) hold. There exists a $C > 0$ depending only on derivatives of y such that

$$\|\varepsilon_n\| \leq Ct_n h^{q+1}, \qquad n \geq 0$$

for all problems (3.1) with $z \in D$, if and only if

$$\sup_{z \in D} \|\varphi(z)\| < \infty.$$

Proof. Assume $\|\varphi(z)\| \leq \gamma$ on D. We obtain from (3.3)

$$\delta_n = \left(I - R(z)\right)\varphi(z)h^{q+1}y^{(q+1)}(t_n) + O(h^{q+2})$$

where the constant involved in the $O(h^{q+2})$-term depends only on $\max_t \|y^{(q+2)}(t)\|$. The order $q + 1$ convergence result now follows easily, for example by writing out the recursion (3.2) in full and using partial summation, similar to formula (15.10) in Hairer & Wanner (1991, Chapter IV).

Now, suppose that $\|\varphi(z_0)\| = \infty$ for some $z_0 \in D$. Then for any $K > 0$ we can select a $z \in \operatorname{int}(D)$ such that $\|\varphi(z)\| > K$. We consider (3.1) with solution $y(t) = t^{q+1}/(q+1)!$. The recursion for the global errors shows that

$$\varepsilon_n = \big(I - R(z)^n\big)\varphi(z)h^{q+1}.$$

Let $h \to 0$ while $z = h\lambda$ and $t_n = t$ are kept fixed. Because $z \in \operatorname{int}(D)$ we know that $R(z)^n \to 0$ as $n \to \infty$. Hence for $h \to 0$ it holds that

$$h^{-q-1}\|\varepsilon_n\| \to \|\varphi(z)\| > K.$$

Since K can be taken arbitrarily large, there is no convergence with order $q + 1$.

□

The above theorem is similar to results obtained in Burrage *et al* (1986) and Dekker *et al* (1986) for Runge–Kutta methods with $D = \mathbb{C}^- = \{z \in \mathbb{C} : \operatorname{Re}(z) \leqslant 0\}$.

It is clear that $\|\varphi(z)\|$ is bounded near $z = z_0$ if $I - R(z_0)$ is non-singular. This holds certainly in the interior of the stability region S. For z_0 on the boundary of S it may happen that φ has a pole at z_0. To have order $q + 1$ convergence such points should be excluded from D. This limits, of course, the class of problems under consideration. Considering applications for systems of equations it is not reasonable to exclude the origin, and therefore we will consider the boundedness of φ near $z = 0$ in more detail. In the next lemma p denotes the classical order of the method.

LEMMA 3.2 Suppose that $p \geqslant q + l$ and $\|(I - R(z))^{-1}\| = O(z^{-l})$, $z \to 0$. Then φ is bounded near $z = 0$.

Proof. If $\lambda = 1$ and $y(t) = t^{q+1}/(q+1)!$, the local error equals

$$\delta_n = h^{q+1}(d_1 + hB_1d_2 + h^2B_1B_2d_2 + h^3B_1B_2^2d_2 + \cdots).$$

A convergence criterion of Skeel (1976) for non-stiff problems gives

$$\delta_n = O(h^p), \qquad E\delta_n = O(h^{p+1})$$

where E is the eigenprojection (or component) of A_1 for eigenvalue 1, see Hairer & Wanner (1991, pp 292, 293). It holds that $Ev = 0$ iff $v \in \operatorname{Range}(I - A_2)$. Hence, if $p \geqslant q + 1$ we have

$$d_1 \in \operatorname{Range}(I - A_1),$$

if $p \geqslant q + 2$ then

$$d_1 = 0, \qquad B_1d_2 \in \operatorname{Range}(I - A_1),$$

and if $p \geqslant q + l$, $l \geqslant 3$ it holds that

$$d_1 = B_1B_2^jd_2 = 0 \quad \text{for} \quad 0 \leqslant j \leqslant l - 3, \qquad B_1B_2^{l-2}d_2 \in \operatorname{Range}(I - A_1).$$

First, suppose that $l = 1$. Then $d_1 = (I - A_1)x$ for some vector x, and we get

$$\begin{aligned}\varphi(z) &= (I - R(z))^{-1}((I - A_1)x + zB_1(I - zB_2)^{-1}d_2)\\ &= x + z(I - R(z))^{-1}B_1(I - zB_2)^{-1}(d_2 + A_2x)\end{aligned}$$

which is bounded near $z = 0$ by the assumption on $(I - R(z))^{-1}$.

The proof for $l = 2, 3, \ldots$ can be obtained by induction, using the following argument. Suppose

$$\varphi(z) = z^j(I - R(z))^{-1}B_1(I - zB_2)^{-1}B_2^{l-1}d_2, \qquad B_1B_2^{l-1}d_2 = (I - A_1)x.$$

Then

$$\begin{aligned}\varphi(z) &= z^j(I - R(z))^{-1}((I - A_1)x + zB_1(I - zB_2)^{-1}B_2^l d_2)\\ &= z^jx - z^{j+1}(I - R(z))^{-1}B_1(I - zB_2)^{-1}(B_2^l d_2 + A_2x). \qquad \square\end{aligned}$$

COROLLARY 3.3 Suppose $p \geq q + 1$ and 1 is a simple eigenvalue of A_1. Then φ is bounded near $z = 0$. Moreover, if D is closed in $\bar{\mathbb{C}}$ and $\rho(R(z)) < 1$ for all $z \in D - \{0\}$ then φ is bounded on D.

Proof. Since 1 is a simple eigenvalue of A_1, there is exactly one eigenvalue $\lambda_j(z)$ of $R(z)$ such that $\lambda_j(z) = 1 + O(z)$. Order p implies

$$\lambda_j(z) = e^z + O(z^{p+1}).$$

It follows that $\det(I - R(z))^{-1} = O(z^{-1})$ as $z \to 0$, and consequently the same holds for $\|(I - R(z))^{-1}\|$. $\square$

3.2 *The stability conditions* (3.4), (3.5)

Suppose that $I - zB_2$ is non-singular for all $z \in S$. Then (3.5) will hold for any bounded set $D \subset S$. Further it is easily seen that if B_2 is non-singular, or if

$$B_1 = (B_{11} \quad B_{12}), \qquad B_2 = \begin{pmatrix} 0 & 0 \\ B_{21} & B_{22} \end{pmatrix}$$

with B_{22} non-singular and $B_{11} - B_{12}B_{22}^{-1}B_{21} = 0$, then $zB_1(I - zB_2)^{-1}$ will also be bounded near $z = \infty$.

To study condition (3.4), let

$$a(z) = \sup_{n \geq 0} \|R(z)^n\|.$$

We shall take D to be compact in $\bar{\mathbb{C}}$. Then for (3.4) to hold it is sufficient that for any $z_0 \in D$ there is a neighbourhood on which $a(z)$ is bounded.

First, consider $z_0 \in \operatorname{int}(S)$. There $\rho(R(z)) < 1$, and so we can find a norm $\|\cdot\|_0$ for which $\|R(z_0)\|_0 < 1$. By continuity of $R(z)$, whose elements are rational functions without poles in S, there exists a neighbourhood U of z_0 and a $\kappa < 1$ such that $\|R(z)\|_0 \leq \kappa$ on U. Therefore

$$\begin{aligned}\|R(z)^n - R(z_0)^n\|_0 &= \left\|\sum_{j=0}^{n-1} R(z)^j(R(z) - R(z_0))R(z_0)^{n-j-1}\right\|_0\\ &\leq n\kappa^{n-1}K_0|z - z_0|\end{aligned}$$

for some $K_0 > 0$. Using equivalence of norms we see that there is a $K > 0$ such that

$$\sup_{n \geq 0} \|R(z)^n - R(z_0)^n\| \leq K\,|z - z_0|$$

for all $z \in U$. So, $a(z)$ is continuous in z_0 and thus certainly bounded near z_0.

Next, consider $z_0 \in S$ on the boundary of S. Let us assume for the moment that the eigenvalues of $R(z_0)$ with modulus 1 are simple. We call these eigenvalues $\lambda_1(z_0), \ldots, \lambda_m(z_0)$. Taking a sufficiently small, closed neighbourhood U of z_0, these eigenvalues remain simple and we can decompose $R(z)$ as

$$R(z) = \lambda_1(z)P_1(z) + \cdots + \lambda_m(z)P_m(z) + Q(z) \tag{3.7}$$

with eigenprojections $P_1(z), \ldots, P_m(z)$, $P_i(z)P_j(z) = \delta_{ij}P_i(z)$ and $\rho(Q(z)) < 1$, $P_i(z)Q(z) = 0$. Such a spectral decomposition can be obtained for example from the Jordan normal form: if

$$R = X\begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_l \end{pmatrix}X^{-1}, \qquad J_1 = \lambda_1 \qquad \text{then} \qquad P_1 = Xe_1e_1^{\mathrm{T}}X^{-1}$$

with e_1 denoting the first unit vector. Since $\lambda_1(z), \ldots, \lambda_m(z)$ are simple on U, the corresponding eigenprojections are continuous (even analytic, see Kato (1982, pp 77–79)) and thus also $Q(z)$ is continuous on U.

Define for $z \in U$

$$\bar{a}(z) = \|P_1(z)\| + \cdots + \|P_m(z)\| + \sup_{n \geq 0} \|Q(z)^n\|.$$

In the same way as before, using $\rho(Q(z)) < 1$, it follows that $\bar{a}(z)$ is continuous on U. Moreover, since

$$R(z)^n = \lambda_1(z)^n P_1(z) + \cdots + \lambda_m(z)^n P_m(z) + Q(z)^n$$

and $|\lambda_j(z)| \leq 1$ on S, we see that $a(z) \leq \bar{a}(z)$ on $U \cap S$. Hence $a(z)$ is bounded on $U \cap S$.

In general, $R(z_0)$ may have multiple eigenvalues of modulus 1, as long as such eigenvalues have a complete set of eigenvectors. Again let $\lambda_1(z), \ldots, \lambda_m(z)$ be the eigenvalues of $R(z)$ for which $|\lambda_j(z)| \to 1$ as $z \to z_0$. We assume that on a neighbourhood U of z_0

$$\lambda_j(z) \quad \text{is simple on} \quad U_0 = U - \{z_0\}, \qquad 1 \leq j \leq M. \tag{3.8a}$$

Then the decomposition (3.7) is valid on U_0 and the previous arguments can be repeated if the $P_j(z)$ are bounded on U_0. We will show that this is ensured under the assumption

$$|\lambda_i(z) - \lambda_j(z)| \geq \gamma d(z, z_0) \qquad \text{if} \qquad z \in U, \qquad 1 \leq i \neq j \leq m \tag{3.8b}$$

with $\gamma > 0$ and $d(z, z_0) = |z - z_0|$ if z_0 is finite, $d(z, z_0) = |z^{-1}|$ in case $z_0 = \infty$.

We consider, for convenience of notation, only $P_1(z)$ near $z_0 = 0$. Let $x_1(z)$ be the eigenvector for $\lambda_1(z)$, with the norm of $x_1(0)$ equal to 1. Since $P_1(z)$ is the projection onto (multiples of) $x_1(z)$ along the other eigen- and principal vectors $x_2(z),\dots,x_k(z)$ of $R(x)$, it follows from a geometrical consideration that $\|P_1(z)\|\to\infty$, $z\to 0$ iff

$$\inf\{\|x_1(z)-v\| : v\in \mathrm{Span}\,(x_2(z),\dots,x_k(z))\}\to 0 \qquad \text{as} \qquad z\to 0.$$

Suppose that $\lambda_1(0)=\cdots=\lambda_r(0)$ with the other eigenvalues having a different limit. Premultiplication of $x_1(z)-v$ with powers of $R(z)-\lambda_i(z)I$, $i\geqslant r+1$, then gives

$$\inf\{\|x_1(z)-v\| : v\in \mathrm{Span}\,(x_2(z),\dots,x_r(z))\}\to 0 \qquad \text{as} \qquad z\to 0. \quad (3.9)$$

It will be shown below that under (3.8) the vectors $x_1(0), x_2(0),\dots,x_r(0)$ are linearly independent. This implies of course that

$$x_1(0)\notin \mathrm{Span}\,(x_2(0),\dots,x_r(0)),$$

but also that $\mathrm{Span}\,(x_2(z),\dots,x_r(z))$ is continuous near $z=0$, and thus we get a contradiction with (3.9).

It remains to prove that

$$a_1x_1(0)+\cdots+a_rx_r(0)=0$$

implies $a_1,\dots,a_r=0$. In view of (3.8) we know that

$$\lambda_j(z)=\mu+zv_j+\mathrm{O}(z), \qquad j=1,\dots,r$$

with $|\mu|=1$ and distinct v_j. For the corresponding eigenvectors we have

$$a_jx_j(z)=w_j+zw_j'+\mathrm{O}(z)$$

and $R(z)$ has an expansion

$$R(z)=R_0+zR_1+\mathrm{O}(z^2).$$

Comparing powers, it follows that

$$(R_0-\mu I)w_j=0, \qquad (R_1-v_jI)w_j=-(R_0-\mu I)w_j', \qquad j=1,\dots,r.$$

Let $w=v_1w_1+\cdots+v_rw_r$. Since $w_1+\cdots+w_r=0$, we obtain

$$(R_0-\mu I)w=0, \qquad w=(R_0-\mu I)(w_1'+\cdots+w_r'). \quad (3.10)$$

However, $R_0=R(z_0)$ is power bounded. Therefore (3.10) implies $w=0$; otherwise there would be a Jordan block for μ with dimension larger than one.

Now, from $w=0$ it follows that $\tilde{w}_1+\cdots+\tilde{w}_{r-1}=0$, $\tilde{w}_j=(v_j-v_r)w_j$. We can repeat the above argument to arrive finally at

$$(v_1-v_2)\cdots(v_1-v_r)w_1=0.$$

Since $w_1=a_1x_1(0)$, $x_1(0)\neq 0$ and the v_j are distinct we must have $a_1=0$. In the same way it follows that the other a_j are zero.

Summarizing the above, we have proved the following result.

LEMMA 3.4 Let $D\subset S$ be closed in $\bar{\mathbb{C}}$, and suppose (3.8) holds for any z_0 on the boundary of D. Then there is an $\alpha>0$ such that (3.4) is valid.

The above result is well known for linear multistep methods. A proof for these methods based on the Kreiss matrix theorem can be found in Hairer & Wanner (1991). This proof can also be used for general linear methods if there are no multiple eigenvalues of modulus 1 on the boundary of D.

The following example shows that (3.8) cannot be replaced by a weaker condition on the eigenvalues.

EXAMPLE Consider

$$\tilde{R}(z)=\begin{pmatrix}\lambda_1(z) & 0\\ \kappa(z) & \lambda_2(z)\end{pmatrix}$$

near $z=0$. Suppose

$$\lambda_j(z)=1+\nu_j z+\mathrm{O}(z^2), \qquad \kappa(z)=\gamma z+\mathrm{O}(z^2)$$

with $\nu_j>0$ and $\gamma\neq 0$. This $\tilde{R}(z)$ may represent a principal submatrix of the Schur normal form of $R(z)$. We have

$$\tilde{R}(z)^n=\begin{pmatrix}\lambda_1(z)^n & 0\\ \sigma_n(z) & \lambda_2(z)^n\end{pmatrix}, \qquad \sigma_n=\kappa(z)\sum_{m=1}^{n}\lambda_1(z)^{n-m}\lambda_2(z)^{m-1}.$$

We consider $z\in\mathbb{C}^-$ such that

$$|e^{nz}|=1+\mathrm{O}(nz^2), \qquad |nz|\to\infty, \qquad |nz^2|\to 0$$

in the limit $n\to\infty$. For instance, z on a circle in the left half plane touching the origin, $|z|\sim n^{-\frac{3}{4}}$ (by the assumption $\nu_j>0$ there is a small circle of this type lying in the stability region of $\tilde{R}(z)$). If $\nu_1=\nu_2$ it follows easily that

$$|\sigma_n(z)|=|\gamma nz|\,(1+\mathrm{O}(nz^2))\to\infty.$$

Note that this may even happen if $\tilde{R}(z)^n$ is power bounded for all $z\in\mathbb{C}^-$ (A-stability).

3.3 *Linear systems*

The formulae derived for the scalar case remain valid for linear systems (1.3) with $z=hL$. Therefore, the results trivially carry over to the case where L is a normal matrix with $h\lambda\in D$ for all eigenvalues $\lambda\in L$.

For arbitrary L satisfying (1.4) we can use the scalar results by setting $D=\mathbb{C}^-$ and employing the inequality

$$\|\Psi(hL)\|\leq\sup_{\zeta\in\mathbb{C}^-}\|\Psi(\zeta)\|$$

which is valid for $h>0$, L satisfying (1.4) and any $\Psi(\zeta)$ which is a rational function or a matrix valued function whose elements are rational. This is a variant of a theorem of J. von Neumann, see Hairer & Wanner (1991), Nevalinna (1985). In a direct way we thus obtain the following corollary, where we denote for brevity by $C(y)$ some positive constant whose value is determined by bounds on the derivatives of y.

COROLLARY 3.5 Consider the class of linear problems (1.3), (1.4) and assume that (3.4), (3.5) hold with $D = \mathbb{C}^-$. We have

$$\|\varepsilon_n\| \leq C(y) t_n h^q, \qquad n \geq 0.$$

Moreover, if $p \geq q + 1$ then

$$\|\varepsilon_n\| \leq C(y) t_n h^{q+1}, \qquad n \geq 0$$

if and only if $\|\varphi(z)\|$ is uniformly bounded on $\mathbb{C}^-$. □

Note that in the second part of this corollary the assumption $p \geq q + 1$ can be removed (since $\|\varphi(0)\| < \infty$ implies already $p \geq q + 1$), but the result would not become more general.

Error bounds of the above type are relevant for initial boundary problems for partial differential equations via the method of lines. Then L will represent a discretized differential operator and g will contain forcing terms and boundary conditons. In fact, in the proof of Theorem 3.1 we have considered $h \to 0$ while at the same time $\lambda \to \infty$, which is natural for partial differential equations if the time and space grid are refined simultaneously and λ is a 'large' eigenvalue of L. Until now we have not made any assumption on g. However, if the initial boundary value problem has homogeneous boundary conditions, then g and its derivatives will be bounded independently of the mesh width in space, see for instance Sanz-Serna & Verwer (1989), Verwer (1986). This is equivalent to saying that derivatives of Ly will be bounded.

Global error bounds based on derivatives of y and Ly can be easily obtained by writing the local error (3.3) in the form

$$\begin{aligned} \delta_n = {} & \frac{1}{z}(d_1 + zB_1(I - zB_2)^{-1} d_2) h^{q+2} L y^{(q+1)}(t_n) \\ & + \sum_{j \geq q+2} (d_1^{(j)} + zB_1(I - zB_2)^{-1} d_2^{(j)}) h^j y^{(j)}(t_n). \end{aligned} \tag{3.11}$$

We denote by $C(y, Ly)$ some constant whose value is determined by bounds on the derivatives of y and Ly.

THEOREM 3.6 Consider the class of linear problems and assume that (3.4), (3.5) hold with $D = \mathbb{C}^-$. If $p \geq q + 1$ it holds that

$$\|\varepsilon_n\| \leq C(y, Ly) t_n h^{q+1}, \qquad n \geq 0.$$

For $p \geq q + 2$ we have

$$\|\varepsilon_n\| \leq C(y, Ly) t_n h^{q+2}, \qquad n \geq 0$$

if and only if $\left\|\frac{1}{z}\varphi(z)\right\|$ is uniformly bounded on $\mathbb{C}^-$.

Proof. We may assume for the proof that $L = \lambda$ is scalar. Afterwards von Neumann's theorem can be applied. Note also that the first part in the theorem is trivial if $d_1 = 0$, since then the local error can be bounded by $C_0(y, Ly) h^{q+2}$. Let us assume only that

$$d_1 = (I - A_1)x.$$

We have

$$\frac{1}{z}(d_1 + zB_1(I - zB_2)^{-1}d_2)h = \frac{1}{\lambda}(I - R(z))x + hv(z)$$

where

$$v(z) = B_1(I - zB_2)^{-1}(d_2 + A_2 x).$$

Hence the leading term in δ_n can be written as

$$(I - R(z))xh^{q+1}y^{(q+1)}(t_n) + v(z)h^{q+2}\lambda y^{(q+1)}(t_n).$$

Convergence with order $q+1$ now follows as before by writing out the recursion for the global errors and using partial summation.

For the remaining order $q+2$ result we observe that we only have to consider the leading term of the local error; the second term can be treated similar to above and so will give an order $q+2$ contribution to the global error. Starting from (3.11) it can now be shown in the same way as in the proof of Theorem 3.1 that the boundedness of $\varphi(z)/z$ is necessary and sufficient for convergence with order $q+2$. □

Similar to Corollary 3.3 it can be shown that $\varphi(z)/z$ will be bounded near $z=0$ if $p \geq q+2$ and 1 is a simple eigenvalue of A_1. So, the order $q+2$ convergence result is valid for all methods which also satisfy $\rho(R(z)) < 1$ for $z \in \mathbb{C}^- - \{0\}$, $z = \infty$.

The fact that the boundedness of $\varphi(z)/z$ is sufficient for convergence with order $q+2$ in the above theorem gives a generalization of a result in Brenner *et al* (1982) for Runge–Kutta methods. On the order hand, the class of equations considered in Brenner *et al* (1982) is more general than (1.3), (1.4), namely with operators L on a Banach space.

It is clear that one can continue in the same way to obtain a higher-order result based on y, Ly and L^2y. An interesting limiting case arises if we require the derivatives of L^jy to be bounded for arbitrary powers of j. This is a reasonable assumption for pure initial value problems for partial differential equations (or for problems where the boundary condition is replaced by a periodicity condition). Then one obtains the classical order of convergence, since all elementary differentials needed in the theory based on Butcher trees will be bounded.

4. Non-linear problems

The methods in this section are supposed to be algebraically stable. This means that there exists a symmetric positive definite G and a diagonal positive semi-definite D such that the matrix

$$M = \begin{pmatrix} G - A_1^T G A & A_2^T D - A_1^T G B_1 \\ DA_2 - B_1^T G A_1 & DB_2 + B_2^T D - B_1^T G B_1 \end{pmatrix} \tag{4.1}$$

is positive semi-definite. The existence of such G, D implies that for arbitrary dissipative problems (1.5) the inequality

$$\|\bar{u}_n - u_n\|_G \leq \|\bar{u}_0 - u_0\|_G$$

is valid for (2.1), (2.2) with ξ_j, $\eta_j = 0$ in the norm generated by G, see Burrage & Butcher (1980), Butcher (1987a), Butcher (1987b), Hairer & Wanner (1991). Thus (2.7) will hold with a constant α determined by G.

Further it is known from the theory of Runge–Kutta methods that (2.8), as well as (2.9) are fulfilled for the dissipative problems if there is a positive definite diagonal matrix $\tilde{D}$ such that

$$\tilde{D}B_2 + B_2^{\mathrm{T}}\tilde{D} \quad \text{is positive definite,} \tag{4.2}$$

see for instance Dekker & Verwer (1984), Frank *et al* (1985a), Hairer & Wanner (1991).

In the following we will consider only general linear methods that are irreducible, in the sense that no internal stages are redundant (similar to the irreducibility concept of Dahlquist and Jeltsch for Runge–Kutta methods, see Dekker & Verwer (1984), Hairer & Wanner (1991) for example).

LEMMA 4.1 For an irreducible, algebraically stable method it holds that the matrix D in (4.1) is positive definite and

$$B_1 v \in \text{Range}\,(I - A_1) \Rightarrow e^{\mathrm{T}} D v = 0$$

for any vector v.

Proof. Let k_0 be the preconsistency vector. The fact that M is positive semi-definite implies

$$e^{\mathrm{T}}D = k_0^{\mathrm{T}} G B_1, \qquad k_0^{\mathrm{T}} G (I - A_1) = 0, \tag{4.3}$$

see Hairer & Wanner (1991, p 386). Therefore, if $B_1 v = (I - A_1)w$ we have

$$e^{\mathrm{T}} dv = k_0^{\mathrm{T}} G B_1 v = k_0^{\mathrm{T}} G (I - A_1) w = 0.$$

Let $e_i = (0, 0, \ldots, 1, \ldots, 0)^{\mathrm{T}}$ be the ith unit vector. It is clear from (4.3) that $De_i = 0$ if $B_1 e_i = 0$. On the other hand, we have for all w

$$2w^{\mathrm{T}} B_2^{\mathrm{T}} D w \geq w^{\mathrm{T}} B_1^{\mathrm{T}} G B_1 w,$$

so that $De_i = 0$ implies $(B_1 e_i)^{\mathrm{T}} G B_1 e_i = 0$, and since G is positive definite we obtain $B_1 e_i = 0$. Thus

$$B_1 e_i = 0 \Leftrightarrow De_i = 0. \tag{4.4}$$

Now, suppose that certain columns of B_1 are zero, say

$$B_1 e_i = 0, \qquad i = 1, \ldots, r, \qquad B_1 e_i \neq 0, \qquad i = r + 1, \ldots, k.$$

The positive semi-definiteness of $DB_2 + B_2^{\mathrm{T}} D$, together with $De_i = 0$ iff $i \leq r$, implies that

$$e_j^{\mathrm{T}} B_2 e_i = 0, \qquad i \leq r, \qquad j > r$$

(this follows easily from an argument similar to Hairer & Wanner (1991, p 200)). However, this means that $v_{n1}, \ldots, v_{nr}$ in (1.2) do not influence the other internal vectors v_{nj} nor the external vectors $u_{n+1,i}$. So the first r internal stages are superfluous. This shows that for an irreducible method we cannot have zero columns in B_1, and consequently $De_i \neq 0$ for all i. □

Again, it is clear from Section 2 that algebraic stability and (4.2) lead to convergence for arbitrary dissipative problems with order of convergence q. As for the linear problems we consider the question whether the order can be $q+1$. The next theorem is a generalization of a result of Burrage & Hundsdorfer (1987) for Runge–Kutta methods.

THEOREM 4.2 Let the method be irreducible, algebraically stable and such that (4.2) is valid. Assume further that $c_i - c_j$ is not an integer if $i \neq j$. Then, there exists a constant $C > 0$ only depending on derivatives of y such that

$$\|\varepsilon_n\| \leqslant C t_n h^{q+1}, \qquad n \geqslant 0$$

for all problems (1.1), (1.5), if and only if

$$d_1 = (I - A_1)x, \qquad d_2 = -A_2 x \tag{4.5}$$

for some vector x.

Proof. Define

$$\Phi(Z) = \big(I - R(Z)\big)^{-1}\big(d_1 + B_1 Z(I - B_2 Z)^{-1} d_2\big)$$

for $Z = \operatorname{diag}(z_1, \ldots, z_s)$. If (4.5) is valid then $\Phi(Z) \equiv x$. Convergence with order $q+1$ follows easily by considering the modified errors

$$\tilde{\varepsilon}_n = \varepsilon_n - x h^{q+1} y^{(q+1)}(t_n) \tag{4.6}$$

which satisfy a recursion

$$\tilde{\varepsilon}_{n+1} = R(Z_n)\tilde{\varepsilon}_n + \tilde{\delta}_n, \qquad \|\tilde{\delta}_n\| \leqslant C_1 h^{q+2}$$

with C_1 only depending on derivatives of y, see for instance Burrage & Hundsdorfer (1987), Hundsdorfer & Steininger (1991).

To prove the necessity of (4.5), consider

$$y'(t) = \lambda(t) y(t) + g(t)$$

with $\operatorname{Re}\big(\lambda(t)\big) \leqslant 0$. This can be converted to a real two-dimensional problem satisfying (1.5). We choose λ to be h-periodic, giving

$$Z_n = Z = \operatorname{diag}(z_1, \ldots, z_s) \qquad \text{with} \qquad z_j = h\lambda(t_n + c_j h).$$

By our assumption on the c_j these z_j can be chosen to be mutually independent and arbitrary in $\mathbb{C}^-$. In the same way as in the proof of Theorem 3.1 it follows that convergence with order $q+1$ only holds if there is a $\gamma > 0$ such that

$$\|\Phi(Z)\| \leqslant \gamma \qquad \text{for all} \qquad Z = \operatorname{diag}(z_1, \ldots, z_s), \qquad \operatorname{Re} z_j \leqslant 0. \tag{4.7}$$

Each component of $\Phi(Z)$ is rational in $z_1, \ldots, z_s$ and we can write, with the same denominator for all components,

$$\Phi_j(Z) = \frac{f_{j0}(Z) + f_{j1}(Z) + f_{j2}(Z) + \cdots}{g_0(Z) + g_1(Z) + g_2(Z) + \cdots}$$

where $g_0(Z)$ is constant, $g_1(Z)$ linear in $z_1,\dots,z_s$, $g_2(Z)$ contains the quadratic terms $z_m z_l$, etc, and likewise for $f_{j0}(Z), f_{j1}(Z),\dots$. Suppose that $g_0,\dots,g_{k-1}\equiv 0$ and $g_k \not\equiv 0$ for some $k \geq 0$. Boundedness of $\Phi(Z)$ for small $z_j \in \mathbb{C}^-$ implies that $f_{j0},\dots,f_{j,k-1}\equiv 0$. Moreover, $f_{jk}(Z)$ has to be a multiple of $g_k(Z)$; otherwise we could choose $z_m = i\tau\sigma_m$ on the imaginary axis with fixed σ_m such that $g_m(Z)=0$ for all τ but $f_{jk}(Z)\neq 0$, and this would give unboundedness of $\Phi_j(Z)$ for $\tau \to 0$. It follows that (4.7) only holds if there is a vector x such that

$$\Phi(Z) = \big(g_k(Z) + \mathrm{O}(\tau^{k+1})\big)^{-1}\big(g_k(Z)x + \mathrm{O}(\tau^{k+1})\big) \qquad \text{for} \qquad |z_m| \leq \tau, \quad 1 \leq m \leq s.$$

From the definition of $\Phi(Z)$ it is seen that

$$(I - A_1)\Phi(Z) - d_1 = B_1 Z(I - B_2 Z)^{-1}\big(A_2\Phi(Z) + d_2\big). \tag{4.8}$$

We now consider $z_m = \tau\rho_m$ with arbitrary but fixed $\rho_m \in \mathbb{C}^-$ such that $g_k(Z)\neq 0$ for $\tau > 0$. Then $g_k(Z) = \delta\tau^k$ with $\delta \neq 0$. Multiplication of (4.8) with the denominator of $\Phi(Z)$ and comparison of powers of τ gives

$$\delta\tau^k\big((I - A_1)x - d_1\big) = \mathrm{O}(\tau^{k+1}), \tag{4.9}$$

and hence $d_1 = (I - A_1)x$. Using Lemma 4.1 it thus follows from (4.8) that

$$e^{\mathrm{T}} DZ(I - B_2 Z)^{-1}\big(A_2\Phi(Z) + d_2\big) = 0,$$

and we obtain in a similar way as (4.9)

$$\delta\tau^k e^{\mathrm{T}} DZ(A_2 x + d_2) = \mathrm{O}(\tau^{k+2}). \tag{4.10}$$

Since the elements of D are all positive and the ρ_j are arbitrary (under the restriction that $g_k(Z)\neq 0$), it follows that $d_2 = -A_2 x$. □

The fact that the modified errors $\tilde{e}_n$ in the above proof have local errors of $\mathrm{O}(h^{q+2})$ can also be interpreted in the following way: the stage order of the method is q with respect to the correct value function $u(t)$, but if we change $u(t)$ to

$$\tilde{u}(t) = u(t) - xh^{q+1}y^{(q+1)}(t)$$

the stage order becomes $q+1$. This is related to the more abstract concept of BH-consistency in Li (1989).

LEMMA 4.3 Suppose the method is irreducible, algebraically stable and 1 is a simple eigenvalue of A_1. Then

$$(4.5) \Rightarrow p = q + 1.$$

Proof. The classical order conditions are necessary for non-stiff dissipative problems, and thus $p \geq q+1$. Let us assume that $p \geq q+2$. Then we know, see Section 3, that

$$d_1 = 0, \qquad B_1 d_2 \in \mathrm{Range}\,(I - A_1).$$

Since $d_1 = (I - A_1)x$ and 1 is a simple eigenvalue, the vector x must be a multiple of the preconsistency vector,

$$x = \delta k_0 \qquad \text{for some} \qquad \delta \neq 0.$$

Hence $d_2 = -\delta e$, and therefore

$$B_1 e \in \text{Range}\,(I - A_1).$$

From Lemma 4.1 it now follows that $e^{\mathrm{T}}De = 0$, which contradicts the fact that D is positive definite. □

It is not known whether high-order algebraically stable methods can fulfil (4.5). The above lemma shows already that for most general linear methods condition (4.5) will not hold. In fact, it is known from Burrage & Hundsdorfer (1987) that there are only very few algebraically stable Runge–Kutta methods with property (4.5) and that the maximal classical order of such methods is 3. There seems to be only one Runge–Kutta method of practical interest for which we obtain convergence with order $q + 1$, namely the implicit midpoint rule, see Kraaijevanger (1985). The implicit midpoint rule can also be regarded as a one-leg multistep method. Most one-leg methods have $p = q + 1$ and for these methods condition (4.5) is always fulfilled, as can be seen from the convergence results in Hundsdorfer & Steininger (1991) and Hairer & Wanner (1991) (or by noting that the functions φ and Φ are identical if $k = 1$), but algebraically stable (G-stable) one-leg multistep methods have $p \geqslant 2$.

Condition (4.5) can be avoided by restricting the class of non-linear problems. Similar to linear problems, along the lines of Burrage *et al* (1985), convergence results with order $q + 1$ can be obtained for semi-linear equations

$$y'(t) = L(t)y(t) + g\big(t, y(t)\big)$$

with a smoothly varying $L(t)$ satisfying (1.4) and with $g(t, y)$ satisfying a Lipschitz condition near the exact solution. This class of problems contains those non-linear dissipative problems for which the partial derivatives

$$\frac{\partial^2}{\partial t\, \partial y} f(t, y) \qquad \text{and} \qquad \frac{\partial^2}{\partial y^2} f(t, y)$$

can be bounded by a moderate constant.

Finally we note that condition (4.5) is also not needed for non-linear singularly perturbed problems, nor is algebraic stability. This can be seen from the results of Hairer *et al* (1988) for Runge–Kutta methods, see also Hairer & Wanner (1991). Convergence results for problems of this type have not yet been derived for general linear methods.

Acknowledgement

This work has been supported by the Fonds National Suisse de la Recherche Scientifique, grant No 20-28923.90.

References

BRENNER, P., CROUZEIX, M., & THOMÉE, V. 1982 Single step methods for inhomogeneous linear differential equations in Banach space. *RAIRO Anal. Numér.* **16,** 5–26.

BURRAGE, K., & BUTCHER, J. C. 1980 Nonlinear stability of a general class of differential equation methods. *BIT* **20,** 185–203.

BURRAGE, K., & CHIPMAN, F. H. 1989 Efficiently implementable multivalue methods for solving stiff ordinary differential equations. *Appl. Numer. Math.* **5,** 23–40.

BURRAGE, K., & HUNDSDORFER, W. H. 1987 The order of B-convergence of algebraically stable Runge–Kutta methods. *BIT* **27,** 62–71.

BURRAGE, K., HUNDSDORFER, W. H., & VERWER, J. G. 1986 A study of B-convergence of Runge–Kutta methods. *Computing* **36,** 17–34.

BUTCHER, J. C. 1987a The equivalence of algebraic and AN-stability. *BIT* **27,** 510–533.

BUTCHER, J. C. 1987b *The Numerical Analysis of Ordinary Differential Equations—Runge–Kutta and General Linear Methods.* Chichester: John Wiley and Sons.

DEKKER, K., KRAAIJEVANGER, J. F. B. M., & SPIJKER, M. N. 1986 The order of B-convergence of the Gaussian Runge–Kutta methods. *Computing* **36,** 35–41.

DEKKER, K., & VERWER, J. G. 1984 *Stability of Runge–Kutta Methods for Stiff Nonlinear Differential Equations* (*CWI Monographs* 2). Amsterdam: North-Holland.

FRANK, R., SCHNEID, J., & UEBERHUBER, C. W. 1985a Stability properties of implicit Runge–Kutta methods. *SIAM J. Numer. Anal.* **22,** 497–514.

FRANK, R., SCHNEID, J., & UEBERHUBER, C. W. 1985b Order results for implicit Runge–Kutta methods applied to stiff systems. *SIAM J. Numer. Anal.* **22,** 515–534.

HARIER, E., LUBICH, CH., & ROCHE, M. 1988 Error of Runge–Kutta methods for stiff problems studied via differential algebraic equations. *BIT* **28,** 678–700.

HAIRER, E., NØRSETT, S. P., & WANNER, G. 1987 *Solving Ordinary Differential Equations I—Nonstiff Problems* (*Springer Series in Computational Mathematics* 8). Berlin: Springer.

HAIRER, E., & WANNER, G. 1991 *Solving Ordinary Differential Equations II—Stiff and Differential-Algebraic Problems* (*Springer Series in Computational Mathematics* 14). Berlin: Springer.

HUNDSDORFER, W. H., & STEININGER, B. I. 1991 Convergence of linear multistep and one-leg methods for stiff nonlinear initial value problems. *BIT* **31,** 124–143.

KATO, T. 1982 *A Short Introduction to Perturbation Theory for Linear Operators.* New-York: Springer.

KRAAIJEVANGER, J. F. B. M. 1985 B-convergence of the implicit midpoint rule and the trapezoidal rule. *BIT* **25,** 652–666.

LI, SHOUFU 1989 Stability and B-convergence of general linear methods. *J. Comput. Appl. Math.* **28,** 281–296.

NEVANLINNA, O. 1985 Matrix valued versions of a result of von Neumann with an application to time discretization. *J. Comput. Appl. Math.* **12–13,** 475–489.

SANZ-SERNA, J. M., & VERWER, J. G. 1989 Stability and convergence at the PDE/stiff ODE interface. *Appl. Numer. Math.* **5,** 117–132.

SKEEL, R. 1976 Analysis of fixed stepsize methods. *SIAM J. Numer. Anal.* **13,** 664–685.

SPIJKER, M. N. 1986 The relevance of algebraic stability in implicit Runge–Kutta methods. *The Numerical Treatement of Differential Equations* (*Teubner Texte zur Math.* 82) (K. Strehmel, ed). Leipzig.

VERWER, J. G. 1986 Convergence and order reduction of diagonally implicit-Runge–Kutta schemes in the method of lines. *Numerical Analysis* (*Pitman Research Notes in Mathematics* 140) (D. F. Griffiths and G. A. Watson, eds). Boston, MA: Pitman.